ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р 59880— 2021

Эргономика. Проектирование и применение испытаний речевых технологий

МЕТОДИКА ОПРЕДЕЛЕНИЯ ПОКАЗАТЕЛЕЙ КАЧЕСТВА СИНТЕЗА РЕЧИ ПО ТЕКСТУ

Издание официальное

Москва
Российский институт стандартизации
2021

gost.e-ecolog.ru

Предисловие

1 РАЗРАБОТАН Федеральным государственным унитарным предприятием «Научно-исследовательский институт «Квант» («ФГУП «НИИ «Квант»)

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 164 «Искусственный интеллект»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 25 ноября 2021 г. № 1583-ст

4 ВВЕДЕН ВПЕРВЫЕ

*Правила применения настоящего стандарта установлены в статье 26 Федерального закона от 29 июня 2015 г. № 162-ФЗ «О стандартизации в Российской Федерации». Информация об изменениях к настоящему стандарту публикуется в ежегодном (по состоянию на 1 января текущего года) информационном указателе «Национальные стандарты», а официальный текст изменений и поправок — в ежемесячном информационном указателе «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ближайшем выпуске ежемесячного информационного указателя «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет (*gost.e-ecolog.ru

© Оформление. ФГБУ «РСТ», 2021

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

Содержание

1 Область применения 1
2 Нормативные ссылки 1
3 Термины и определения 2
4 Приборы и оборудование 3
5 Общие требования 3
6 Измерение смысловой разборчивости синтезированной речи аудитивным методом 4
7 Измерение смысловой разборчивости ускоренной синтезированной речи аудитивным методом 6
8 Измерение интонационной разборчивости синтезированной речи аудитивным методом 6
9 Измерение интонационной разборчивости ускоренной синтезированной речи аудитивным методом 7
10 Измерение естественности синтезированной речи аудитивным методом 7
11 Измерение качества нормализации синтезируемого текста 8
12 Измерение качества управления синтезом речи разметкой в формате SSML 9
13 Неоцениваемые функциональные возможности синтезатора 10
Приложение А (обязательное) Правила составления таблиц для измерения смысловой разборчивости речи 11
Приложение Б (обязательное) Правила составления таблиц для оценки интонационной разборчивости речи 14
Приложение В (обязательное) Правила составления таблиц для измерения естественности речи . . . 20
Приложение Г (обязательное) Правила составления таблиц для проверки нормализации текста 21
Приложение Д (обязательное) Правила составления таблиц для проверки управления синтезом 23
Библиография 25

Введение

Настоящий стандарт регламентирует методику определения показателей качества синтеза речи по тексту, основанную на методах измерения показателей разборчивости и естественности речи. Методика включает в себя измерение разборчивости для нормального и ускоренного темпа речи, показателя деградации разборчивости синтезированной речи, генерируемой синтезатором, и показателя естественности речи и признаков искажения естественности речи (картавость, гнусавость, помеха и др.). Дополнительно данная методика учитывает качество нормализации входного текста (преобразование чисел в числительные, расшифровка сокращений, правильность определения ударений в словах, в том числе разрешение омонимии, правильность интонационного оформления текста), а также качество управления процессом синтеза речи с помощью тегов по стандарту Speech Synthesis Markup Language.

Оценки, установленные настоящим стандартом, могут быть получены в обычном (ручном) и автоматизированном (программном) режимах. В программном режиме использована персональная электронно-вычислительная машина (ПЭВМ), что обеспечивает автоматизацию измерительных процедур (аудирование, подсчет результатов измерения, поиск аномальных измерений, оформление протоколов измерений и пр.).

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

Эргономика.
Проектирование и применение испытаний речевых технологий

МЕТОДИКА ОПРЕДЕЛЕНИЯ ПОКАЗАТЕЛЕЙ КАЧЕСТВА СИНТЕЗА РЕЧИ ПО ТЕКСТУ

Ergonomics. Design and application of speech technology tests. Methodology for determining the quality indicators of speech synthesis from text

Дата введения — 2022—03—01

1 Область применения

1.1 Настоящий стандарт распространяется на синтез речи по тексту [синтезатор речи, преобразования «текст в речь» («text to speech»)] и устанавливает показатели качества синтеза речи и методы измерений:

- смысловой разборчивости синтезированной речи с нормальным и ускоренным темпами методом аудитивных измерений;
- интонационной разборчивости синтезированной речи с нормальным и ускоренным темпами методом аудитивных измерений;
- естественности синтезированной речи методом аудитивных измерений величины искажений естественности речи;
- качества управления синтезом речи разметкой в формате Speech Synthesis Markup Language (SSML) согласно [1];
- качества нормализации текста.

1.2 Настоящий стандарт распространяется только на синтезаторы речи по тексту класса «Информатор» и «Собеседник» в соответствии с 3.10.

2 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты:

ГОСТ 13107 Устройства приема и передачи речи. Методы измерения акустических шумов

ГОСТ 15150 Машины, приборы и другие технические изделия. Исполнения для различных климатических районов. Категории, условия эксплуатации, хранения и транспортирования в части воздействия климатических факторов внешней среды

ГОСТ Р 7.0.12—2011 Библиографическая запись. Сокращение слов и словосочетаний на русском языке. Общие требования и правила

ГОСТ Р 53188.1 Государственная система обеспечения единства измерений. Шумомеры. Часть 1. Технические требования

Примечание — При пользовании настоящим стандартом целесообразно проверить действие ссылочных стандартов в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет или по ежегодному информационному указателю «Национальные стандарты», который опубликован по состоянию на 1 января текущего года, и по выпускам ежемесячного информационного указателя «Национальные стандарты» за текущий год. Если заменен ссылочный стандарт, на который дана недатированная ссылка, то рекомендуется использовать версию этого стандарта с учетом всех внесенных в данную версию изменений. Если заменен ссылочный стандарт, на который дана датиро-

Издание официальное

ванная ссылка, то рекомендуется использовать версию этого стандарта с указанным выше годом утверждения (принятия). Если после утверждения настоящего стандарта в ссылочный стандарт, на который дана датированная ссылка, внесено изменение, затрагивающее положение, на которое дана ссылка, то это положение рекомендуется применять без учета данного изменения. Если ссылочный стандарт отменен без замены, то положение, в котором дана ссылка на него, рекомендуется применять в части, не затрагивающей эту ссылку.

3 Термины и определения

В настоящем стандарте применены следующие термины с соответствующими определениями:

3.1

разборчивость синтезированной речи: Относительное количество (в процентах) правильно синтезированных элементов (слогов, слов, фраз) артикуляционных таблиц.

[ГОСТ Р 50840—95, раздел 3]

Примечание — В настоящем стандарте с целью повышения точности оценки разборчивость синтезированной речи подразделена на два типа: смысловую разборчивость речи и интонационную разборчивость речи.

3.2 **смысловая разборчивость речи:** Относительное количество синтезированных фраз по заданному тексту с учетом правильной постановки ударений в словах, в том числе с учетом разрешения случаев омографии и правильности автоматической замены буквы «е» на «ё».

3.3 **интонационная разборчивость речи:** Относительное количество правильно синтезированных фраз по заданному тексту с учетом интонационного оформления.

Примечание — Показатель качества интонационной разборчивости не применим для синтезаторов речи класса «Информатор» (3.10).

3.4 **естественность синтезированной речи:** Величина, характеризующая субъективную оценку соответствия звучания синтезированной речи естественному произношению.

3.5 **нормальный темп синтезированной речи:** Синтез речи по тексту со скоростью приблизительно от 8 до 12 звуков в секунду.

Примечание — В целях удобства измерения скорость можно измерять в буквах в секунду. Для русского языка это будет также от 8 до 12 букв в секунду. Точность измерения в буквах в секунду является достаточной для определения класса темпа речи.

3.6 **ускоренный темп синтезированной речи:** Синтез речи по тексту со скоростью примерно 20 букв в секунду.

3.7 **нормализация текста:** Преобразование встречающихся в тексте чисел, аббревиатур, сокращений, символов, дат, времени, номеров телефонов, символов валют, иноязычных слов в соответствии с литературной нормой произнесения.

3.8 **интонационное оформление синтезированной речи:** Реализация в синтезированной речи интонации высказываний, соответствующей знакам препинания в тексте.

Примечание — Предложение, заканчивающееся точкой, должно оформляться интонацией завершения; предложение, заканчивающееся знаком вопроса, — вопросительной интонацией; предложение, заканчивающееся восклицательным знаком, — восклицательной интонацией; предложение, заканчивающееся многоточием или двоеточием, — интонацией незавершенности. Интонация незавершенности также должна быть реализована и в том случае, когда в тексте имеется запятая. В этом случае перед паузой (запятой) должна быть интонация незавершенности.

3.9 **омография:** Раздел лингвистики, изучающий одинаковые по написанию слова, имеющие различие в произношении в зависимости от контекста.

3.10 **классы синтезаторов речи:** Разделение синтезаторов речи по функциональности выражения интонации речи.

Примечание — В настоящем стандарте все синтезаторы речи разделены на 3 класса:

- «Информатор» — класс интонационной функциональности, который предполагает синтез речи для текста, содержащего только повествовательные предложения или вопросительные предложения;
- «Собеседник» — класс интонационной функциональности предполагает, что текст может содержать все виды предложений, включая предложения с интонацией незавершенности, интонация имеет нейтральную эмоциональную окраску.

- «Актер» — класс интонационной функциональности, который обеспечивает синтез любых видов интонации и эмоциональной окраски.

Уровень функциональности по способу выражения интонации речи увеличивается от класса «Информатор» к классу «Актер». При увеличении уровня функциональности возрастает количество критериев качества. Настоящий стандарт определяет критерии качества только для классов «Информатор» и «Собеседник».

4 Приборы и оборудование

4.1 При выполнении измерений должны быть применены измерительные системы, испытательное оборудование и средства измерений по ГОСТ Р 53188.1, ГОСТ 15150:

- программное обеспечение, выполняющее синтез речи по заданному тексту;
- динамические головные телефоны;
- шумомер по ГОСТ Р 53188.1, класс точности 3.

4.2 Для выполнения измерений допускается использование дополнительного оборудования:

- ПЭВМ;
- внешний носитель с тестами для проведения измерений.

5 Общие требования

5.1 Измерения следует проводить в нормальных климатических условиях по ГОСТ 15150.

5.2 Уровень акустического шума в помещении для испытаний должен быть не более 50 дБ (по шкале А ГОСТ 13107).

5.3 Прослушивание должно быть выполнено на головные телефоны, обеспечивающие звуковое давление не выше 79 дБ.

5.4 Измерения проводит бригада аудиторов. Аудитор должен быть в возрасте от 18 до 50 лет, не иметь явных дефектов слуха.

5.5 Размер бригады аудиторов зависит от измеряемого показателя качества.

5.6 Объем, представительность и вариативность тестового материала (число таблиц и их состав) должны быть такими, чтобы обеспечить полноту проверки объекта измерений и исключить возможность предварительной настройки синтезатора на конкретный состав тестового материала.

5.7 При повторном измерении показателя качества аудитор должен прослушивать тот тестовый материал, который он ранее не прослушивал на этапе обучения или в предыдущих измерениях.

5.8 Продолжительность одной непрерывной сессии измерений, производимой аудиторами, не должна превышать 45 мин. В сутки рекомендуется проводить не более 4 сессий, с перерывами между ними не менее 20 мин.

5.9 Результаты измерений для всех методов должны быть оформлены протоколом, содержащим следующую информацию:

- объект испытаний;
- результаты измерений.

При выполнении измерений естественности протокол должен дополнительно содержать следующую информацию:

- ширина спектра аудиосигнала: узкополосный (от 300 до 3700 Гц) или широкополосный (от 50 до 7000 Гц);
- средняя длительность синтезируемой речи: кратковременная (одно предложение, не более 10 слов), долговременная (три или более предложений, связанных общим смыслом, 20 или более слов);
- среднее значение оценки естественности для аудиозаписей, содержащих естественную речь;
- средние значения оценки естественности для аудиозаписей базовых синтезаторов (при наличии).

Рекомендуется дополнительно предоставлять следующую информацию:

- цель испытаний;
- дату и место проведения испытаний;
- условия и методику проведения испытаний;
- сведения о головных телефонах, использовавшихся для воспроизведения аудиосигнала;
- количество аудиторов;
- количество и длительность сеансов;
- профили участников, распределение по возрасту и полу;
- типы задействованных участников, например неопытные участники или эксперты.

6 Измерение смысловой разборчивости синтезированной речи аудитивным методом

6.1 Измерения проводит бригада аудиторов в количестве не менее 15 человек. Состав бригады — произвольный.

6.2 Бригада аудиторов работает в два этапа.

6.3 На первом этапе проводят обучение аудиторов на подготовленном речевом тесте. Речевой тест должен содержать мужской и женский голоса, естественную речь, синтезированную речь со всеми категориями ошибок. Аудиторы тренируются определять категорию ошибки синтезированной речи по образцам.

6.4 На втором этапе выполняют измерения испытуемого синтезатора. Синтезатор генерирует речь для таблицы А.1 (приложение А) и примеров, приведенных в А.2 (приложение А), с постоянным уровнем громкости в нормальном темпе речи.

6.5 Пауза между фразами при прослушивании должна составлять (3 ± 2) с. Длительность паузы в заданном интервале определяется аудитором. Уровень громкости устанавливается аудитором на испытательном предложении «Не видали мы такого невода».

6.6 Аудитор прослушивает синтезированное предложение и ставит оценку разборчивости в диапазоне от 1 до 5. Оценку 1 ставят при фиксировании ошибок категории 4. Описание категорий ошибок указано в таблице 1. Оценку 2 ставят, когда наблюдаются ошибки категории 3. Оценку 3 — ошибки категории 2. Оценку 4 — ошибки категории 1. Оценку 5 ставят, когда в синтезированной речи отсутствуют ошибки. При наблюдении ошибок из разных категорий выбирают ошибку с наибольшей категорией.

При измерении смысловой разборчивости следует помнить, что при этом не учитывается точность реализации интонационного оформления фразы. Необходимо принимать во внимание только возможные ошибки, перечисленные в таблице 1.

Таблица 1 — Категории ошибок в оценке смысловой разборчивости синтезированной речи

Категория	Описание
4	Пропадание и полное искажение слов
3	Частичное искажение слов
2	Неправильные ударения в словах
1	Неправильное положение пауз
1	Неправильная длительность пауз
1	Неестественность звучания речи
1	Вставка шума в речевых паузах

6.7 Измерение разборчивости включает в себя прослушивание всеми аудиторами всех фраз таблиц, подготовленных отдельно для мужского и женского голосов. Для каждого измерения выбирают разные таблицы. Протокол измерения смысловой разборчивости речи должен содержать дату измерения, идентификатора аудитора, идентификатор голоса синтезатора, идентификатор фразы и оценку разборчивости.

6.8 Проводят единичные измерения как средние значения разборчивости по оценкам аудитора для каждой пары таблица — голос. Для каждого измерения вычисляют среднее значение разборчивости S согласно [1] по формуле

$$S = \frac{1}{N}\sum_{i=1}^{N} S_i, \qquad (1)$$

где N — число единичных измерений (таблиц);

S_i — единичное измерение (среднее значение в таблице 1).

Далее выявляют сомнительные значения S_i, которые отбрасывают, и вычисляют новое значение S. Методика обработки приведена в 6.14.

6.9 Аудиторы, результаты которых имеют более двух отклонений от средних значений по бригаде аудиторов более чем на величины, указанные в таблице 2, подлежат замене или исключению из бригады.

Таблица 2 — Максимальное нормальное отклонение

Среднее значение разборчивости по бригаде, по шкале от 1 до 5	Отклонение от среднего значения
4,55 и более	0,05
4,30—4,54	0,06
4,05—4,29	0,07
3,01—4,04	0,08
3,00 и менее	0,09

6.10 Измерения считают законченными при достижении бригадой повторяемых результатов измерения. Повторяемость результатов проверяют по статистической значимости различий средних величин разборчивости по *t*-критерию Стьюдента с уровнем значимости 0,05. Итоговым измерением является объединение последних единичных измерений, не имеющих статистически значимых различий.

6.11 После прослушивания в течение 45 мин делают перерыв на 20 мин. Время работы бригады должно быть не более 4 ч в сутки.

6.12 Классы разборчивости и нормы разборчивости речи — в соответствии с таблицей 3.

Таблица 3 — Характеристики класса разборчивости

Класс разборчивости	Характеристика класса разборчивости	Норма разборчивости речи для синтеза речи по тексту по шкале от 1 до 5
1	Понимание синтезированной речи без напряжения внимания. Ошибки произнесения отсутствуют	>4,65
2	Понимание синтезируемой речи без напряжения внимания, редкие ошибки произнесения. Ошибки категории 1	4,30—4,64
3	Понимание синтезируемой речи с некоторым напряжением внимания, без повторений. Ошибки категории 2	3,80—4,29
4	Понимание синтезированной речи с напряжением внимания, с повторениями. Ошибки категории 3	3,05—3,79
5	Частичное понимание синтезированной речи, с повторениями. Ошибки категории 4	1,00—3,04

6.13 Метод следует использовать при аттестации синтезатора речи по тексту, не требующему нормализации.

6.14 Методика обработки данных измерений

С целью исключения спорных результатов измерений показателя качества производят следующую обработку полученных данных:

а) вычисляют среднее значение разборчивости по формуле (1);

б) вычисляют среднее квадратичное отклонение (СКО) по формуле

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(S_i - S)^2}{N-1}}; \qquad (2)$$

в) единичные измерения, для которых $|S_i - S| > 3\sigma$, исключают и производят вычисление нового среднего значения по формуле

$$S = \frac{1}{N-k}\sum_{i=1}^{N-k} S_i, \quad (3)$$

где N — число единичных измерений;

k — число исключенных измерений.

7 Измерение смысловой разборчивости ускоренной синтезированной речи аудитивным методом

7.1 Измерения выполняют по методике раздела 6, за исключением установки темпа речи, приведенной в 6.4. Устанавливают ускоренный темп речи, как правило, от 20 знаков/с. Если необходимо измерение при более высоком ускорении, то это должно быть указано при организации аттестации синтезатора. В результате выполнения этой методики получают оценку разборчивости ускоренной синтезированной речи S_y.

7.2 Коэффициент деградации смысловой разборчивости ускоренной синтезированной речи рассчитывают по формуле

$$D_S = \frac{S_y}{S_n}, \quad (4)$$

где S_n — оценка разборчивости, полученная при выполнении методики по разделу 6.

7.3 Метод следует использовать при аттестации синтезатора речи по тексту, применяемого в режиме синтеза ускоренной речи.

8 Измерение интонационной разборчивости синтезированной речи аудитивным методом

8.1 Измерения проводит бригада аудиторов в количестве не менее 15 человек. Состав бригады произвольный.

8.2 Бригада аудиторов работает в два этапа.

8.3 На первом этапе проводят обучение аудиторов на подготовленном речевом тесте. Речевой тест должен содержать мужской и женский голоса, естественную речь, синтезированную речь со всеми категориями ошибок. Аудиторы тренируются определять категорию ошибки синтезированной речи по образцам.

8.4 На втором этапе выполняют измерения испытуемого синтезатора. Синтезатор генерирует речь для таблицы Б.1 (приложение Б) с постоянным уровнем громкости в нормальном темпе речи.

8.5 Пауза между фразами при прослушивании должна составлять (3 ± 2) с. Длительность паузы в заданном интервале определяется аудитором. Уровень громкости устанавливается аудитором на испытательном предложении «Не видали мы такого невода».

8.6 Аудитор прослушивает синтезированное предложение и ставит оценку разборчивости интонации — 0 или 1. Оценка 1 означает, что интонация точно соответствует знакам препинания; оценка 0 — интонация не соответствует. Монотонность интонации не должна влиять на оценку разборчивости.

8.7 Измерение интонационной разборчивости включает в себя прослушивание всеми аудиторами всех фраз подготовленной таблицы. Протокол измерения интонационной разборчивости речи должен содержать дату измерения, идентификатор аудитора, идентификатор голоса синтезатора, идентификатор фразы и оценку разборчивости.

8.8 Оценку интонационной разборчивости речи S, %, вычисляют по формуле

$$S = 100 \cdot \left(\frac{1}{N}\sum_{i=1}^{N} S_i \right), \quad (5)$$

где N — число фраз в подготовленной таблице;

S_i — значение оценки интонации отдельной фразы.

9 Измерение интонационной разборчивости ускоренной синтезированной речи аудитивным методом

9.1 Измерения проводят по методике раздела 8, за исключением установки темпа речи, приведенной в 8.4. Устанавливают ускоренный темп речи, как правило, от 20 знаков/с. Если необходимо измерение при более высоком ускорении, то это должно быть указано при организации аттестации синтезатора. По данной методике выполняют оценку разборчивости ускоренной синтезированной речи S_y.

9.2 Коэффициент деградации интонационной разборчивости ускоренной синтезированной речи D_S вычисляют по формуле:

$$D_S = \frac{S_y}{S_n}, \quad (6)$$

где S_n — оценка разборчивости, полученная при выполнении методики по разделу 8.

9.3 Метод следует использовать при аттестации синтезатора речи по тексту, используемого в режиме синтеза ускоренной речи.

10 Измерение естественности синтезированной речи аудитивным методом

10.1 Измерения проводит бригада аудиторов в количестве не менее 20 человек. Отклонение доли мужчин от доли женщин не должно превышать 20 %. Размер бригады определен для уровня значимости 0,05 с доверительным интервалом оценки среднего значения естественности, равным 0,03 балла по [1].

10.2 Бригада аудиторов работает в два этапа.

10.3 На первом этапе проводят обучение аудиторов на подготовленном речевом тесте. Речевой тест должен содержать мужской и женский голоса, естественную речь, синтезированную речь со всеми уровнями естественности. Аудиторы тренируются оценивать естественность синтезированной речи в баллах по образцам.

10.4 На втором этапе выполняют измерения испытуемого синтезатора. Синтезатор генерирует речь в нормальном темпе речи для таблиц, составленных в соответствии с правилами приложения В. Формируется тест для аудиторов как случайная последовательность различных аудиозаписей, содержащих либо естественную речь, либо синтезированную. Расстояние между аудиозаписями с естественной речью в последовательности аудиозаписей должно быть не более 5.

10.5 Пауза между фразами при прослушивании должна составлять (3 ± 2) с. Длительность паузы в заданном интервале определяется аудитором. Уровень громкости устанавливается аудитором на испытательном предложении «Не видали мы такого невода».

10.6 Аудитор прослушивает аудиозаписи, ставит оценку естественности звучания речи по степени искажений речи. Субъективную оценку проводят в абсолютных категориях по 5-балльной шкале в соответствии с таблицей 4. Прилагательное «абсолютный» в наименовании шкалы указывает на то, что аудиторам необходимо оценить каждую аудиозапись безотносительно к другим аудиозаписям.

Таблица 4 — Шкала оценок естественности речи

Оценка естественности звучания речи	Значение, балл	Степень искажений речи
Отлично	5	Искажения отсутствуют. Синтезированная речь не отличается от естественной
Хорошо	4	Искажения присутствуют только в отдельных словах. Каждое слово разборчиво. Возможна монотонная интонация
Удовлетворительно	3	Искажения присутствуют постоянно
Плохо	2	Искажения присутствуют постоянно. Некоторые слова неразборчивы
Неприемлемо	1	Пропадание и полное искажение слов

10.7 Измерение естественности включает в себя прослушивание аудиторами всех аудиозаписей не менее пяти таблиц, подготовленных отдельно для мужского и женского голосов. Для каждого измерения выбирают разные таблицы. Протокол измерения естественности речи должен содержать дату измерения, идентификатор аудитора, идентификатор голоса синтезатора или диктора, ширину спектра аудиосигнала, идентификатор фразы и оценку естественности.

10.8 Вычисляют единичные измерения как средние значения естественности по оценкам аудитора для каждой пары таблицы и голоса. Для каждого измерения вычисляют среднее значение естественности Q по формуле

$$Q = \frac{1}{N}\sum_{i=1}^{N} Q_i, \tag{7}$$

где N — число единичных измерений;

Q_i — единичное измерение.

Далее выявляют спорные значения Q_i, которые не учитывают, и вычисляют новое значение Q. Методика обработки приведена в 6.14.

10.9 Измерения считают законченными при достижении бригадой повторяемых результатов измерения. Повторяемость результатов проверяют по статистической значимости различий средних величин естественности по t-критерию Стьюдента с уровнем значимости 0,05. Итоговым измерением является объединение последних единичных измерений, не имеющих статистически значимых различий.

10.10 После прослушивания 10 таблиц делают перерыв на 20 мин. Время работы бригады должно быть не более 4 ч за один день. Общее число таблиц за один рабочий день до 40.

10.11 Расчет оценок естественности выполняют отдельно по различным синтезированным голосам и совместно по всем естественным голосам.

10.12 Расчет оценки естественности синтезатора речи в целом определяют как среднее значение оценок естественности по отдельным голосам синтезатора.

10.13 Оценку естественности используют для сравнения синтезаторов речи с учетом ширины спектра аудиосигнала и средней длительности синтезируемой речи. Среднее значение естественности для синтезатора с узкополосным аудиосигналом уменьшается на 0,2 при сравнении со средним значением синтезатора с широкополосным аудиосигналом. Оценка естественности уменьшается с увеличением средней длительности синтезируемой речи.

10.14 Сравнивать оценки естественности синтезаторов речи, выполненных в двух различных измерениях, возможно только после корректировки различий начал отсчета координат и масштабов субъективных шкал экспериментов. Базой коррекции шкал являются оценки естественности для естественной речи и оценки естественности для базовых синтезаторов. Базовые синтезаторы — это синтезаторы, которые оценивают в нескольких измерениях.

11 Измерение качества нормализации синтезируемого текста

11.1 Измерения проводит бригада аудиторов в количестве не менее 3 человек.

11.2 Бригада аудиторов работает без тренировки.

11.3 Синтезатор генерирует речь для всех примеров приложения Г с постоянным уровнем громкости в нормальном темпе.

11.4 Пауза между фразами при прослушивании должна составлять (3 ± 2) с. Длительность паузы в заданном интервале определяется аудитором. Уровень громкости устанавливается аудитором на испытательном предложении «Не видали мы такого невода».

11.5 Аудитор прослушивает синтезированную речь, проверяя содержание речи по нормализованной форме текста, и фиксирует число ошибок (количество слов, которые не соответствуют расшифровкам), которые допустил синтезатор. Необходимо учитывать, что в одном предложении может быть несколько случаев нормализации, следовательно, и ошибок может быть несколько. Если ошибки нормализации отсутствуют, то указывают значение 0 (ноль). Протокол измерения качества нормализации синтезируемого текста должен содержать дату измерения, идентификатор аудитора, идентификатор голоса синтезатора, идентификатор фразы, количество случаев нормализации и количество ошибок.

11.6 Формируется итоговый протокол по измерениям аудиторов. Для каждой фразы количество ошибок нормализации принимают как медианное значение количества ошибок, вычисленное по всем аудиторам.

11.7 Качество нормализации синтезируемого текста S_N, %, вычисляют по формуле

$$S_N = 100\left(1 - \frac{N_o}{N}\right), \quad (8)$$

где N_o — общее число ошибок, зафиксированных в таблицах для проверки нормализации;

N — общее число случаев, где требуется нормализация.

12 Измерение качества управления синтезом речи разметкой в формате SSML

12.1 Измерения проводит бригада аудиторов в количестве не менее 3 человек.

12.2 Измерение проводят на рекомендуемом подмножестве SSML-тегов и значений атрибутов [2], приведенном в таблице 5. В выделенное подмножество попали наиболее частотные, используемые на практике теги.

Таблица 5 — Рекомендуемое подмножество SSML-тегов и значений атрибутов

Тег	Атрибут, значение
«break»	strength=(«none», «x-weak», «weak», «medium», «strong», «x-strong») Значения указаны в порядке увеличения длительности паузы. Значение по умолчанию — «medium».
«break»	time=[«0ms», «2s»]
«emphasis»	level=(«none», «reduced», «moderate», «strong») Значение по умолчанию — «moderate».
«p»	---
«prosody»	pitch=[«-50%», «200%»]
	pitch=(«x-low», «low», «medium», «high», «x-high»)
	rate=[«-50%», «200%»]
	rate=(«x-slow», «slow», «medium», «fast», «x-fast»)
	volume=[«-100%», «200%»]
	volume=(«silent», «x-soft», «soft», «medium», «loud», «x-loud»)
«s»	---
«say-as»	stress=«номер гласной»
«say-as»	interpret-as=«date» format=(«mdy», «dmy», «ymd», «md», «ym», «my», «m», «d», «y»)
«say-as»	interpret-as=«time» format=(«hms24», «hms12»)
«say-as»	interpret-as=(«telephone», «characters», «cardinal», «ordinal»)
«sub»	---

12.3 Бригада аудиторов работает в два этапа.

12.4 На первом этапе проводят обучение аудиторов на подготовленном речевом тесте. Речевой тест должен содержать фразы с вариантами правильного и неправильного управления просодией речи. Аудиторы тренируются определять ошибки управления просодией речи по образцам.

12.5 На втором этапе выполняют проверку испытуемого синтезатора. Синтезатор генерирует речь для всех таблиц приложения Д с постоянным уровнем громкости в нормальном темпе, исключая случай управления громкостью и темпом речи.

12.6 Пауза между фразами при прослушивании должна составлять от 1 до 2 с. Длительность паузы в заданном интервале определяется аудитором. Уровень громкости устанавливается аудитором на испытательном предложении «Не видали мы такого невода».

12.7 Аудитор прослушивает синтезированную речь с учетом управляющего действия конкретного SSML-тега и фиксирует число ошибок (число несоответствий требуемому эффекту управляющего воздействия SSML-тега), которые допустил синтезатор. Если ошибки управления отсутствуют, то указывают значение 0 (ноль). Измерение качества управления следует проводить с учетом разборчивости синтезированного текста. Если управляющее воздействие тега привело к нарушению разборчивости произносимых слов, то такое управляющее воздействие считают ошибочным. Протокол измерения качества управления синтезом должен содержать дату измерения, идентификатор аудитора, идентификатор голоса синтезатора, идентификатор фразы и количество ошибок.

12.8 Формируют итоговый протокол по измерениям аудиторов. Для каждой фразы количество ошибок управления принимают в качестве медианного значения количества ошибок, вычисленного по всем аудиторам.

12.9 Качество управления синтезом S_C, %, вычисляют по формуле

$$S_C = 100\left(1 - \frac{N_o}{N}\right), \quad (9)$$

где N_o — общее количество ошибок итогового протокола;

N — общее число фраз, при котором использованы SSML-теги.

12.10 Оценку качества управления синтезированной речью используют для сравнения синтезаторов речи по критерию эффективности управления процессом синтеза.

13 Неоцениваемые функциональные возможности синтезатора

К неоцениваемым функциональным возможностям синтезатора речи методами измерения, которые описаны в настоящем стандарте, можно отнести:

- передачу эмоциональной окраски синтезированной речи. Возможность управления эмоциональной окраской синтезированной речи;
- возможность переноса эмоциональной окраски и стиля речи с фразы образца речи на синтезированный голос;
- возможность настройки идентичности голоса синтезированной речи по голосу целевого диктора с использованием короткого образца речи (длительность не более 1 мин);
- возможность синтезирования несловной речи;
- возможность воспроизведения в синтезированной речи стихотворных текстов;
- пение.

Приложение А
(обязательное)

Правила составления таблиц для измерения смысловой разборчивости речи

А.1 Таблицы для измерения смысловой разборчивости речи должны содержать по 50 уникальных нормализованных предложений. Предложение должно содержать от 1 до 12 слов. Часть предложений (рекомендуется не менее 20) должны содержать в своем составе запятую и/или омограф. В каждой таблице должно быть несколько слов с замененной «ё» на «е».

Измерение смысловой разборчивости производится без учета интонационной разборчивости.

Пример составления такой таблицы с учетом особенностей русского языка представлен в таблице А.1, которая содержит фонетически представительный текст с дополнительными предложениями, содержащими омографы.

Общий объем текста в подобных таблицах должен обеспечивать фонетическую представительность текстов русской речи. Рекомендуемый объем — 10 таблиц по 50 предложений. В этот объем должна быть включена проверка правильности разрешения омографов. Примеры предложений с омографами приведен ниже в этом приложении. Состав предложений с омографами также должен быть дополнен различными вариантами, чтобы избежать возможность предварительной настройки синтезатора.

Таблица А.1 — Пример таблицы для измерения смысловой разборчивости речи

№ п/п	Предложение
1.	В бухту с моря налетел ветерок
2.	Дно у реки хорошее
3.	Мальчик побежал к лагерю
4.	Сигнал тревоги поднял отряд
5.	Прошло, наверно, всего несколько минут
6.	Штурман просил продолжать разворот
7.	Ледяная вода сводила руки
8.	Лошадь тихонько пофыркивала
9.	Часовой стоял на берегу, осматривался
10.	Яркая луна освещала льды
11.	В доме стало холодно
12.	В печке дотлевали последние угли
13.	Солнце медленно опускалось за море
14.	Лошадь неслась вскачь, если так было надо всаднику
15.	Дети остались сиротами
16.	Снег сыпал в лицо
17.	Маяк был уже близко, а видимость не улучшалась
18.	Ветер стучал в ставни
19.	На столе мигала лампа
20.	Летают две стрекоз**ы**
21.	Она стала тихой и молчаливой
22.	Около дома навалены сугробы, а где лопаты
23.	Громко залаяли собаки
24.	Песня лилась плавно и заунывно

Окончание таблицы А.1

№ п/п	Предложение
25.	Командир корабля одобрил решение
26.	Толпа шла за санями, глубоко увязая в снегу
27.	Над домами качались столбы дыма
28.	На берегу было пусто
29.	Начиналась подготовка к весне
30.	Все идет по плану
31.	Стояли жгучие морозы
32.	Собака этой породы будет настоящим другом хозяину
33.	На подоконнике лежала стопка газет
34.	Видишь, скоро начнется небольшой дождик
35.	Уже одежда его высохла, а настроение поднялось
36.	Это была хорошая книга
37.	Высоко неслись тучи
38.	Они долго, но молча курили
39.	Сани глубоко провалились в сугроб
40.	Вдали виднелся маяк
41.	Из окон домов выглядывали рыбаки
42.	Ртутные пары крайне опасны
43.	Дул сырой ветер, клубились тучи
44.	Море глухо клокотало
45.	На берегу стояли люди
46.	Птицы подлетали к болоту
47.	Кот улегся в ногах, мурлыкая
48.	Он чувствовал себя виноватым
49.	Широко распахнулась дверь
50.	В небе плыла стая лебедей

А.2 Примеры различных вариантов омографов, разнообразные аналоги которых необходимо использовать в таблицах русского языка [2], выделенных подчеркиванием:

1. В душе нет мыла.
2. В душе нет покоя.
3. Летают стрекозы.
4. Летают две стрекозы.
5. Я ехал в вагоне метро с букетом красных гвоздик в руке.
6. Я вбил в стену маленький гвоздик.
7. Верной дорогой идете!
8. Дорогой друг, спасибо за добрые слова.
9. И говорить об этом не стоит.
10. За этим человеком стоит вся страна.
11. Моя хата с краю.
12. Мне довелось побывать на краю света.

13. Сухие и полусухие вина хорошо сочетаются с овощами.
14. Вина подсудимого в совершении преступления установлена.
15. Мы живем в совершенно другом мире.
16. Собака этой породы будет настоящим другом хозяину.
17. Наши специалисты ответят на все ваши вопросы.
18. Все прошло по плану.
19. Забудьте обо всем, что вы знали раньше.
20. Всем туристам будет интересна информация об отдыхе в Финляндии.
21. Берете ли вы деньги в долг?
22. Она была в красном берете и черном пальто.
23. На первых порах промысловая деятельность велась с использованием одного арендованного судна.
24. Всем известно, что если пожелтели белки глаз, если колет в правом боку, значит, печень не в порядке.
25. В пищу идут разные части растений: плоды, корни, луковицы.
26. Ртутные пары крайне опасны.
27. Наша компания также имеет возможность купить колесные пары.
28. Белки в живом организме постоянно расщепляются на исходные аминокислоты.
29. Самарское общество охотников конского бега устроило ипподром.
30. Они обратили внимание на старинную бедуинскую забаву — верблюжьи бега.
31. Ворот рубашки должен плотно прилегать к шее.
32. Особенностью сдвижных автоматических ворот является отсутствие направляющих над проемом.
33. Глотка также функционирует как резонатор для формирующихся в гортани звуков.
34. Удовольствие от вина получаешь, когда за столом отодвигаешь момент глотка.
35. Сегодня турецкий морской паром столкнулся с украинским судном.
36. Двигатель с внешним подводом теплоты имеет устройство нагрева и охлаждения из двух подкамер, наполненных жидкостью и паром.
37. Кто сочувствует нашему горю, помогите, чем можете.
38. Я тону без воды и горю без огня.
39. В покровской тюрьме была пресечена подготовка к побегу.
40. Я побегу только комфортные для себя гонки.
41. Сотрудники английского торгового представительства организовали в Гамбурге клуб любителей гребли.
42. В Германии мы гребли в составе восьмерки, финишировав вторыми, уступив лишь восьмерке итальянцев.
43. До сих пор в области регистрировались единичные случаи заражения лишаем.
44. Мы лишаем себя всемогущества, сознательно ограничивая свои возможности.
45. Вот и закончилась эта зимняя сессия, и снова я умудрился закончить ее с одним трояком.
46. Я строю дом своими руками.
47. Мы снова в строю.

Приложение Б
(обязательное)

Правила составления таблиц для оценки интонационной разборчивости речи

Оценку точности интонационной разборчивости предлагается проводить с помощью таблицы, содержащей от 50 и более уникальных нормализованных предложений. Предложение может содержать от 1 до 12 слов. Каждое из этих предложений должно быть повторено 4 раза и различаться только знаком препинания в конце предложения: точка, вопросительный знак, восклицательный знак, многоточие (или двоеточие).

Пример составления такой таблицы с учетом особенностей русского языка представлен в таблице Б.1, которая содержит фонетически представительный текст с дополнительными предложениями, содержащими омографы.

Общий объем текста в подобных таблицах должен обеспечивать фонетическую представительность текстов русской речи. Рекомендуемый объем — 2 таблицы по 50 предложений. В этот объем должна быть включена проверка правильности разрешения омографов. Состав предложений с омографами также должен быть дополнен различными вариантами для того, чтобы избежать возможность предварительной настройки синтезатора.

Таблица Б.1 — Оценка точности интонационной разборчивости

№ п/п	Предложение	Оценка (0/1)
1.	В бухту с моря налетел ветерок.	
2.	В бухту с моря налетел ветерок?	
3.	В бухту с моря налетел ветерок!	
4.	В бухту с моря налетел ветерок…	
5.	Дно у реки хорошее.	
6.	Дно у реки хорошее?	
7.	Дно у реки хорошее!	
8.	Дно у реки хорошее...	
9.	Мальчик побежал к лагерю.	
10.	Мальчик побежал к лагерю?	
11.	Мальчик побежал к лагерю!	
12.	Мальчик побежал к лагерю...	
13.	Сигнал тревоги поднял отряд.	
14.	Сигнал тревоги поднял отряд?	
15.	Сигнал тревоги поднял отряд!	
16.	Сигнал тревоги поднял отряд…	
17.	Прошло, наверно, всего несколько минут.	
18.	Прошло, наверно, всего несколько минут?	
19.	Прошло, наверно, всего несколько минут!	
20.	Прошло, наверно, всего несколько минут...	
21.	Штурман просил продолжать разворот.	
22.	Штурман просил продолжать разворот?	
23.	Штурман просил продолжать разворот!	
24.	Штурман просил продолжать разворот...	
25.	Ледяная вода сводила руки.	

Продолжение таблицы Б.1

№ п/п	Предложение	Оценка (0/1)
26.	Ледяная вода сводила руки?	
27.	Ледяная вода сводила руки!	
28.	Ледяная вода сводила руки...	
29.	Лошадь тихонько пофыркивала.	
30.	Лошадь тихонько пофыркивала?	
31.	Лошадь тихонько пофыркивала!	
32.	Лошадь тихонько пофыркивала...	
33.	Часовой стоял на берегу, осматривался.	
34.	Часовой стоял на берегу, осматривался?	
35.	Часовой стоял на берегу, осматривался!	
36.	Часовой стоял на берегу, осматривался...	
37.	Яркая луна освещала льды.	
38.	Яркая луна освещала льды?	
39.	Яркая луна освещала льды!	
40.	Яркая луна освещала льды...	
41.	В доме стало холодно.	
42.	В доме стало холодно?	
43.	В доме стало холодно!	
44.	В доме стало холодно...	
45.	В печке дотлевали последние угли.	
46.	В печке дотлевали последние угли?	
47.	В печке дотлевали последние угли!	
48.	В печке дотлевали последние угли...	
49.	Солнце медленно опускалось за море.	
50.	Солнце медленно опускалось за море?	
51.	Солнце медленно опускалось за море!	
52.	Солнце медленно опускалось за море...	
53.	Лошадь неслась вскачь, если так было надо всаднику.	
54.	Лошадь неслась вскачь, если так было надо всаднику?	
55.	Лошадь неслась вскачь, если так было надо всаднику!	
56.	Лошадь неслась вскачь, если так было надо всаднику...	
57.	Дети остались сиротами.	
58.	Дети остались сиротами?	
59.	Дети остались сиротами!	
60.	Дети остались сиротами...	

Продолжение таблицы Б.1

№ п/п	Предложение	Оценка (0/1)
61.	Снег сыпал в лицо.	
62.	Снег сыпал в лицо?	
63.	Снег сыпал в лицо!	
64.	Снег сыпал в лицо...	
65.	Маяк был уже близко, а видимость не улучшалась.	
66.	Маяк был уже близко, а видимость не улучшалась?	
67.	Маяк был уже близко, а видимость не улучшалась!	
68.	Маяк был уже близко, а видимость не улучшалась...	
69.	Ветер стучал в ставни.	
70.	Ветер стучал в ставни?	
71.	Ветер стучал в ставни!	
72.	Ветер стучал в ставни...	
73.	На столе мигала лампа.	
74.	На столе мигала лампа?	
75.	На столе мигала лампа!	
76.	На столе мигала лампа...	
77.	Летают две стрекозы.	
78.	Летают две стрекозы?	
79.	Летают две стрекозы!	
80.	Летают две стрекозы...	
81.	Она стала тихой и молчаливой.	
82.	Она стала тихой и молчаливой?	
83.	Она стала тихой и молчаливой!	
84.	Она стала тихой и молчаливой...	
85.	Около дома навалены сугробы, а где лопаты.	
86.	Около дома навалены сугробы, а где лопаты?	
87.	Около дома навалены сугробы, а где лопаты!	
88.	Около дома навалены сугробы, а где лопаты...	
89.	Громко залаяли собаки.	
90.	Громко залаяли собаки?	
91.	Громко залаяли собаки!	
92.	Громко залаяли собаки...	
93.	Песня лилась плавно и заунывно.	
94.	Песня лилась плавно и заунывно?	
95.	Песня лилась плавно и заунывно!	

Продолжение таблицы Б.1

№ п/п	Предложение	Оценка (0/1)
96.	Песня лилась плавно и заунывно...	
97.	Командир корабля одобрил решение.	
98.	Командир корабля одобрил решение?	
99.	Командир корабля одобрил решение!	
100.	Командир корабля одобрил решение...	
101.	Толпа шла за санями, глубоко увязая в снегу.	
102.	Толпа шла за санями, глубоко увязая в снегу?	
103.	Толпа шла за санями, глубоко увязая в снегу!	
104.	Толпа шла за санями, глубоко увязая в снегу...	
105.	Над домами качались столбы дыма.	
106.	Над домами качались столбы дыма?	
107.	Над домами качались столбы дыма!	
108.	Над домами качались столбы дыма...	
109.	На берегу было пусто, точно.	
110.	На берегу было пусто, точно?	
111.	На берегу было пусто, точно!	
112.	На берегу было пусто, точно...	
113.	Начиналась подготовка к весне.	
114.	Начиналась подготовка к весне?	
115.	Начиналась подготовка к весне!	
116.	Начиналась подготовка к весне...	
117.	Все идет по плану.	
118.	Все идет по плану?	
119.	Все идет по плану!	
120.	Все идет по плану...	
121.	Стояли жгучие морозы.	
122.	Стояли жгучие морозы?	
123.	Стояли жгучие морозы!	
124.	Стояли жгучие морозы...	
125.	Собака этой породы будет настоящим другом хозяину.	
126.	Собака этой породы будет настоящим другом хозяину?	
127.	Собака этой породы будет настоящим другом хозяину!	
128.	Собака этой породы будет настоящим другом хозяину...	
129.	На подоконнике лежала стопка газет,	
130.	На подоконнике лежала стопка газет?	

Продолжение таблицы Б.1

№ п/п	Предложение	Оценка (0/1)
131.	На подоконнике лежала стопка газет!	
132.	На подоконнике лежала стопка газет...	
133.	Видишь, скоро начнется небольшой дождик.	
134.	Видишь, скоро начнется небольшой дождик?	
135.	Видишь, скоро начнется небольшой дождик!	
136.	Видишь, скоро начнется небольшой дождик...	
137.	Уже одежда его высохла, а настроение поднялось.	
138.	Уже одежда его высохла, а настроение поднялось?	
139.	Уже одежда его высохла, а настроение поднялось!	
140.	Уже одежда его высохла, а настроение поднялось...	
141.	Это была хорошая книга.	
142.	Это была хорошая книга?	
143.	Это была хорошая книга!	
144.	Это была хорошая книга...	
145.	Высоко неслись тучи.	
146.	Высоко неслись тучи?	
147.	Высоко неслись тучи!	
148.	Высоко неслись тучи...	
149.	Они долго, но молча курили.	
150.	Они долго, но молча курили?	
151.	Они долго, но молча курили!	
152.	Они долго, но молча курили...	
153.	Сани глубоко провалились в сугроб.	
154.	Сани глубоко провалились в сугроб?	
155.	Сани глубоко провалились в сугроб!	
156.	Сани глубоко провалились в сугроб...	
157.	Вдали виднелся маяк.	
158.	Вдали виднелся маяк?	
159.	Вдали виднелся маяк!	
160.	Вдали виднелся маяк...	
161.	Из окон домов выглядывали рыбаки.	
162.	Из окон домов выглядывали рыбаки?	
163.	Из окон домов выглядывали рыбаки!	
164.	Из окон домов выглядывали рыбаки...	
165.	Ртутные пары крайне опасны.	

Окончание таблицы Б.1

№ п/п	Предложение	Оценка (0/1)
166.	Ртутные пары крайне опасны?	
167.	Ртутные пары крайне опасны!	
168.	Ртутные пары крайне опасны...	
169.	Дул сырой ветер, клубились тучи.	
170.	Дул сырой ветер, клубились тучи?	
171.	Дул сырой ветер, клубились тучи!	
172.	Дул сырой ветер, клубились тучи...	
173.	Море глухо клокотало.	
174.	Море глухо клокотало?	
175.	Море глухо клокотало!	
176.	Море глухо клокотало.	
177.	На берегу стояли люди.	
178.	На берегу стояли люди?	
179.	На берегу стояли люди!	
180.	На берегу стояли люди...	
181.	Птицы подлетали к болоту.	
182.	Птицы подлетали к болоту?	
183.	Птицы подлетали к болоту!	
184.	Птицы подлетали к болоту...	
185.	Кот улегся в ногах, мурлыкая.	
186.	Кот улегся в ногах, мурлыкая?	
187.	Кот улегся в ногах, мурлыкая!	
188.	Кот улегся в ногах, мурлыкая...	
189.	Он чувствовал себя виноватым.	
190.	Он чувствовал себя виноватым?	
191.	Он чувствовал себя виноватым!	
192.	Он чувствовал себя виноватым...	
193.	Широко распахнулась дверь.	
194.	Широко распахнулась дверь?	
195.	Широко распахнулась дверь!	
196.	Широко распахнулась дверь...	
197.	В небе плыла стая лебедей.	
198.	В небе плыла стая лебедей?	
199.	В небе плыла стая лебедей!	
200.	В небе плыла стая лебедей...	

Приложение В
(обязательное)

Правила составления таблиц для измерения естественности речи

Таблицы для измерения естественности речи должны быть составлены в зависимости от средней длительности синтезируемой речи:

- для кратковременного синтеза таблицы должны содержать 100 строк. В каждой строке должно быть одно нормализованное предложение. Размер предложения — не более 10 слов. В каждой таблице необходимо обеспечить интонационную представительность предложений, т. е. должны быть утвердительные предложения с точкой в конце и вопросительные. В каждой из этих групп должно быть одно или несколько предложений, содержащих запятую, омограф. В каждой таблице должно быть несколько слов с замененной «ё» на «е»;

- для долговременного синтеза таблицы должны содержать 50 строк. В каждой строке должен быть текст, содержащий не менее трех нормализованных предложений, связанных между собой общим смыслом. Средний размер текста должен быть не менее 20 слов. В каждой таблице необходимо обеспечить интонационную представительность предложений, т. е. должны быть утвердительные повествовательные, вопросительные, восклицательные предложения и предложения, заканчивающиеся многоточием. В каждой из этих групп должно быть одно или несколько предложений, содержащих запятую, омограф. В каждой таблице должно быть несколько слов с замененной «ё» на «е».

Общее число таблиц, представительность и вариативность текстов должны быть такими, чтобы исключить возможность предварительной настройки синтезатора на конкретный состав предложений в таблицах.

Приложение Г
(обязательное)

Правила составления таблиц для проверки нормализации текста

Г.1 Таблицы для проверки нормализации текста должны содержать предложения для проверки нормализации чисел (римских чисел от 0 до 40), сокращений, специальных знаков, иноязычных слов. В каждой строке таблицы по одному предложению. Предложение должно содержать от трех слов и более. Для каждого типа нормализации (числа, сокращения, специальные знаки, иноязычные слова) должна быть отдельная таблица. Количество строк (предложений) в каждой таблице должно быть таким, чтобы охватить общеупотребительные варианты с вариативной словарной представительностью. Рекомендуется составить для каждого варианта нормализации от пяти вариантов примеров и более. Допускается использование нескольких примеров написания чисел в одном предложении.

Число строк в таблицах по разным типам нормализации и вариативность текста (чисел, сокращений, специальных знаков, иноязычных слов) должны быть такими, чтобы исключить возможность предварительной настройки синтезатора на конкретные тексты.

Г.2 Проверка качества нормализации чисел

Проверка качества нормализации чисел должна включать примеры реализации следующих правил:

- примеры использования количественных числительных;
- примеры использования порядковых числительных;
- примеры использования чисел с окончаниями;
- примеры использования чисел со словами (50-летие);
- примеры использования различных склонений чисел;
- примеры использования целых, дробных, отрицательных чисел;
- примеры использования римских цифр.

Г.2.1 Примеры предложений для проверки нормализации чисел

5 и 6 октября 2008 года в Москве был дождь.

Недавно мы отпраздновали 65-летие Победы.

С 80-х годов прошлого века это явление получило массовый характер.

Было опрошено 187 человек, в т. ч. 91 женщина, 67 мужчин и 29 подростков.

В 1660 году король Людовик XIV женился на инфанте Марии-Терезии Австрийской.

Г.2.2 Расшифровки предложений для проверки качества нормализации чисел

Расшифровки чисел даны в скобках после каждой числовой последовательности и выделены жирным шрифтом в нижеприведенных предложениях.

5 (**пятого**) и 6 (**шестого**) октября 2008 (**две тысячи восьмого**) года в Москве был дождь.

Недавно мы отпраздновали 65-летие (**шестидесятипятилетие**) Победы.

С 80-х (**восьмидесятых**) годов прошлого века это явление получило массовый характер.

Было опрошено 187 (**сто восемьдесят семь**) человек, в т. ч. 91 (**девятносто одна**) женщина, 67 (**шестьдесят семь**) мужчин и 29 (**двадцать девять**) подростков.

В 1660 (**тысяча шестьсот шестидесятом**) году король Людовик XIV (**четырнадцатый**) женился на инфанте Марии-Терезии Австрийской.

Г.3 Проверка качества нормализации сокращений

Проверка качества нормализации сокращений должна включать примеры сокращений, образованных по правилам, перечисленным в пункте 4 ГОСТ Р 7.0.12—2011. Для подготовки примеров предложений с сокращениями, также следует использовать материал по ГОСТ Р 7.0.12—2011 (приложение А).

Г.3.1 Примеры предложений для проверки нормализации сокращений

Ул. Бармалеева, д. 12, кв. 36.

К 2020 г. работа будет закончена.

Стихотворение сопровождает комментарий (на с. 275).

Пробка вылетает из бутылки шампанского со скоростью 13 м/с (около 50 км/ч).

Здание было отреставрировано в XX в.

Г.3.2 Расшифровки предложений для проверки нормализации сокращений

Расшифровки сокращений даны в скобках после каждого сокращения и выделены жирным шрифтом.

Ул. (**улица**) Бармалеева, д. (**дом**) 12, кв. (**квартира**) 36.

К 2020 г. (**году**) работа будет закончена.

Стихотворение сопровождает комментарий (на с. (**странице**) 275).

Пробка вылетает из бутылки шампанского со скоростью 13 м/с (**метров в секунду**) [около 50 км/ч (**километров в час**)].

Здание было отреставрировано в XX в. (**веке**).

Г.4 Проверка качества нормализации специальных знаков

Проверка качества нормализации специальных знаков должна включать примеры использования:
- знаков наиболее частотных в использовании валют;
- записи температуры;
- записи процентов;
- записи частотных служебных символов: *, #, ~ и др.

Г.4.1 Примеры предложений для проверки нормализации специальных знаков

Железная дорога стоимостью $4,2 млрд.
Температура воздуха +15 °С.
Так считают 58 % граждан в возрасте от 20 до 35 лет.
Нужно набрать на клавиатуре *#06#.
Есть опасения, что закроются около 40 % предприятий.

Г.4.2 Расшифровки предложений для проверки нормализации специальных знаков

Расшифровки специальных знаков даны в скобках после каждого специального знака и выделены жирным шрифтом.

Железная дорога стоимостью $4,2 млрд (**долларов**).
Температура воздуха + (**плюс**) 15 °С (**градусов Цельсия**).
Так считают 58 % (**процентов**) граждан в возрасте от 20 до 35 лет.
Нужно набрать на клавиатуре: * (**звездочка**) # (**решетка**) 06# (**решетка**).
Есть опасения, что закроются около 40 % (**процентов**) предприятий.

Г.5 Проверка качества нормализации иноязычных слов

Проверка качества нормализации иноязычных слов должна включать следующие примеры использования:
- наименований интернет адресов;
- наименований компаний, торговых марок, марок автомобилей и другой техники;
- наименований информационных агентств;
- наименований стран, континентов, морей, рек, городов и других часто используемых географических наименований;
- имен и фамилий известных людей;
- частотных аббревиатур.

Г.5.1 Примеры предложений для проверки нормализации иноязычных слов

Наш адрес в Интернете: gost.e-ecolog.ru
Аукционный дом Sotheby's открыл выставку в Третьяковской галерее.
Реклама размещена на каналах CNN и Euronews.
Магазины IKEA начали летнюю распродажу.
Следует установить программу на ваш PC.

Г.5.2 Расшифровки предложений для проверки нормализации иноязычных слов

Расшифровки иноязычных вставок даны в скобках после каждой иноязычной вставки и выделены жирным шрифтом в нижеприведенных предложениях.

Наш адрес в Интернете: gost.e-ecolog.ru **три даблйю точка спичпро точка ру**).
Аукционный дом Sotheby's (**Сотбис**) открыл выставку в Третьяковской галерее.
Реклама размещена на каналах CNN (**СиЭнЭн**) и Euronews (**Евроньюс**).
Магазины IKEA (**Икеа**) начали летнюю распродажу.
Следует установить программу на ваш PC (**ПиСи**).

Приложение Д
(обязательное)

Правила составления таблиц для проверки управления синтезом

Таблицы для проверки управления синтезом должны содержать строки текста, содержащие варианты управляющих сочетаний: SSML-тегов/атрибутов. В каждой строке таблицы должны быть от одного до нескольких предложений, различающихся только сочетаниями тег/атрибут тега. Состав и правила применения тегов с атрибутами в тексте должны соответствовать стандарту SSML. Пример таблицы представлен в таблице Д.1.

Выбор атрибутов тегов должен быть таким, чтобы явным образом показать работоспособность тега. Это означает, что значения атрибутов (для тегов break, prosody, emphasis) должны представлять среднее значение и возможный диапазон на границах, при которых действие тега происходит требуемым образом, но при этом речь все еще остается разборчивой. Оценка разборчивости требует прослушивания нескольких предложений для каждого сочетания тег/атрибут тега (это относится к таким тегам, как break, prosody, emphasis).

Некоторые теги не имеют атрибутов, например теги , <s>. Эти теги в таблицах применяют без атрибутов.

Таблица Д.1 — Примеры применения тегов с атрибутами

Предложение, включающее тег и его атрибут
В тишине слышалось <break strength=«x-weak»/> его дыхание.
В тишине слышалось <break strength=«weak»/> его дыхание.
В тишине слышалось <break strength=«none»/> его дыхание.
В тишине слышалось <break strength=«strong»/> его дыхание.
В тишине слышалось <break strength=«x-strong»/> его дыхание.
Я возьму это. <break time=«0ms»/> Вы не против?
Я возьму это. <break time=«200ms»/> Вы не против?
Я возьму это. <break time=«1000ms»/> Вы не против?
Я возьму это. <break time=«2s»/> Вы не против?
Это <emphasis> большая </emphasis> машина!
Это <emphasis level=«none»> большая </emphasis> машина!
Это <emphasis level=«strong»> большая </emphasis> машина!
Это <emphasis level=«reduced»> большая </emphasis> машина!
Подними голову и посмотри вокруг.
<prosody volume=«-75%»> Доброе утро! </prosody>
<prosody volume=«0%»> Доброе утро! </prosody>
<prosody volume=«200%»> Доброе утро! </prosody>
<prosody volume=«silent»> Доброе утро! </prosody>
<prosody volume=«x-soft»> Доброе утро! </prosody>
<prosody volume=«soft»> Доброе утро! </prosody>
<prosody volume=«medium»> Доброе утро! </prosody>
<prosody volume=«loud»> Доброе утро! </prosody>
<prosody volume=«x-loud»> Доброе утро! </prosody>
<prosody rate=«-50%»> Темп речи важен. </prosody>
<prosody rate=«0%»> Темп речи важен. </prosody>
<prosody rate=«100%»> Высокий темп речи важен для рекламы. </prosody>

Окончание таблицы Д.1

Предложение, включающее тег и его атрибут
<prosody rate=«200%»> Очень высокий темп речи важен для слепых. </prosody>
<prosody rate=«x-slow»> Темп речи важен. </prosody>
<prosody rate=«slow»> Темп речи важен. </prosody>
<prosody rate=«medium»> Темп речи важен. </prosody>
<prosody rate=«fast»> Темп речи важен. </prosody>
<prosody rate=«x-fast»> Очень высокий темп речи важен для слепых </prosody>
<prosody pitch=«-50%»> Высота голоса регулируется. </prosody>
<prosody pitch=«0%»> Высота голоса регулируется. </prosody>
<prosody pitch=«50%»> Высота голоса регулируется. </prosody>
<prosody pitch=«100%»> Высота голоса регулируется. </prosody>
<prosody pitch=«x-low»> Высота голоса регулируется. </prosody>
<prosody pitch=«low»> Высота голоса регулируется. </prosody>
<prosody pitch=«medium»> Высота голоса регулируется. </prosody>
<prosody pitch=«high»> Высота голоса регулируется. </prosody>
<prosody pitch=«x-high»> Высота голоса регулируется. </prosody>
<s>Это пример одного предложения </s> А это другое предложение.
Всем <say-as literal=«здрасте»>здравствуйте</say-as> и не болеть.
<say-as stress=«2»>Здравствуйте</say-as>.
<say-as stress=«3»>Здравствуйте</say-as>.
Сейчас <say-as interpret-as=«date» format=«mdy»>3/6/22</say-as>.
Сейчас <say-as interpret-as=«date» format=«dmy»>6/3/22</say-as>.
Сейчас <say-as interpret-as=«date» format=«ymd»>22/3/6</say-as>.
Сейчас <say-as interpret-as=«date» format=«md»>3/6</say-as>.
Сейчас <say-as interpret-as=«date» format=«ym»>22/3</say-as>.
Сейчас <say-as interpret-as=«date» format=«my»>3/22</say-as>.
Сейчас <say-as interpret-as=«date» format=«m»>3</say-as>.
Сейчас <say-as interpret-as=«date» format=«d»>6</say-as>.
Сейчас <say-as interpret-as=«date» format=«y»>22</say-as>.
Точное время <say-as interpret-as=«time» format=«hms24»> 12:34:56 </say-as>.
Точное время <say-as interpret-as=«time» format=«hms12»> 14:00:00 </say-as>.
Номер телефона <say-as interpret-as=«telephone»>2222230</say-as>.
<say-as interpret-as=«characters»> Д5и3-PWq*7R </say-as>.
<say-as interpret-as=«cardinal» format=«.» detail=«,»> 222,333.22 </say-as>.
<say-as interpret-as=«ordinal»>2230</say-as>.
_{МИД}.

Библиография

[1] Шварц Г. «Выборочный метод. Руководство по применению статистических методов оценивания». — Москва, 1978

[2] Speech Synthesis Markup Language (SSML) Version 1.1 (gost.e-ecolog.ru)

УДК 004.896:006.354 ОКС 35.240.99

Ключевые слова: эргономика, испытания речевых технологий, качество синтеза речи по тексту

Редактор *Л.С. Зимилова*
Технический редактор *В.Н. Прусакова*
Корректор *С.В. Смирнова*
Компьютерная верстка *Е.О. Асташина*

Сдано в набор 29.11.2021. Подписано в печать 29.12.2021. Формат 60×84⅛. Гарнитура Ариал.
Усл. печ. л. 3,72. Уч.-изд. л. 3,35.

Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

Создано в единичном исполнении в ФГБУ «РСТ»
для комплектования Федерального информационного фонда стандартов,
117418 Москва, Нахимовский пр-т, д. 31, к. 2.
gost.e-ecolog.ru gost.e-ecolog.ru